Sub a1

~~LOCAL AUTHENTICATION IN A COMMUNICATION SYSTEM~~

BACKGROUND

I. Field of the Invention

The present invention relates to communication systems, more particularly, to local authentication of a communication system subscriber.

II. Background

The field of wireless communications has many applications including, e.g., cordless telephones, paging, wireless local loops, personal digital assistants (PDAs), Internet telephony, and satellite communication systems. A particularly important application is cellular telephone systems for mobile subscribers. (As used herein, the term "cellular" systems encompasses both cellular and personal communications services (PCS) frequencies.) Various over-the-air interfaces have been developed for such cellular telephone systems including, e.g., frequency division multiple access (FDMA), time division multiple access (TDMA), and code division multiple access (CDMA). In connection therewith, various domestic and international standards have been established including, e.g., Advanced Mobile Phone Service (AMPS), Global System for Mobile (GSM), and Interim Standard 95 (IS-95). In particular, IS-95 and its derivatives, IS-95A, IS-95B, ANSI J-STD-008 (often referred to collectively herein as IS-95), and proposed high-data-rate systems for data, etc. are promulgated by the Telecommunication Industry Association (TIA) and other well known standards bodies.

Cellular telephone systems configured in accordance with the use of the IS-95 standard employ CDMA signal processing techniques to provide highly efficient and robust cellular telephone service. Exemplary cellular telephone systems configured substantially in accordance with the use of the IS-95 standard are described in U.S. Patent Nos. 5,103,459 and 4,901,307, which are

assigned to the assignee of the present invention and fully incorporated herein by reference. An exemplary described system utilizing CDMA techniques is the cdma2000 ITU-R Radio Transmission Technology (RTT) Candidate Submission (referred to herein as cdma2000), issued by the TIA. The standard for cdma2000 is given in draft versions of IS-2000 and has been approved by the TIA. The cdma2000 proposal is compatible with IS-95 systems in many ways. Another CDMA standard is the W-CDMA standard, as embodied in 3rd Generation Partnership Project "3GPP", Document Nos. 3G TS 25.211, 3G TS 25.212, 3G TS 25.213, and 3G TS 25.214.

Given the ubiquitous proliferation of telecommunications services in most parts of the world and the increased mobility of the general populace, it is desirable to provide communication services to a subscriber while he or she is travelling outside the range of the subscriber's home system. One method of satisfying this need is the use of an identification token, such as the Subscriber Identity Module (SIM) in GSM systems, wherein a subscriber is assigned a SIM card that can be inserted into a GSM phone. The SIM card carries information that is used to identify the billing information of the party inserting the SIM card into a mobile phone. Next generation SIM cards have been renamed as USIM (UTMS SIM) cards. In a CDMA system, the identification token is referred to as a Removable User Interface Module (R-UIM) and accomplishes the same purpose. Use of such an identification token allows a subscriber to travel without his or her personal mobile phone, which may be configured to operated on frequencies that are not used in the visited environment, and to use a locally available mobile phone without incurring costs in establishing a new account.

Although convenient, the use of such identification tokens to access account information of a subscriber can be insecure. Currently, such identification tokens are programmed to transmit private information, such as a cryptographic key used for message encryption or an authentication key for identifying the subscriber, to the mobile phone. A person contemplating the theft of account information can accomplish his or her goal by programming a mobile phone to retain private information after the identification token has been

removed, or to transmit the private information to another storage unit during the legitimate use of the mobile phone. Mobile phones that have been tampered in this manner will hereafter be referred to as "rogue shells." Hence, there is a current need to preserve the security of the private information stored on an identification token while still facilitating the use of said private information to access communication services.

Summary

A novel method and apparatus for providing secure authentication to a subscriber roaming outside his or her home system is presented. In one aspect, a subscriber identification token is configured to provide authentication support to a mobile unit, wherein the mobile unit conveys information to the subscriber identification token for transformation via a secret key.

In another aspect, a subscriber identification module for providing local authentication of a subscriber in a communication system is presented, comprising: a memory; and a processor configured to implement a set of instructions stored in the memory, the set of instructions for: generating a plurality of keys in response to a received challenge; generating an authentication signal based on a received signal and a first key from the plurality of keys, wherein the received signal is transmitted from a communications unit communicatively coupled to the subscriber identification module, and the received signal is generated by the communications unit using a second key from the plurality of keys, the second key having been communicated from the subscriber identification module to the communications unit; and transmitting the authentication signal to the communications system via the communications unit.

In another aspect, a subscriber identification module is presented, comprising: a key generation element; and a signature generator configured to receive a secret key from the key generation element and information from a mobile unit, and further configured to output a signature to the mobile unit.

In another aspect, an apparatus for providing secure local authorization of a subscriber in a communication system is presented, comprising a subscriber

identification module configured to interact with a communications unit, wherein the subscriber identification module comprises: a key generator for generating a plurality of keys from a received value and a secret value, wherein at least one communication key from the plurality of keys is delivered to the communications unit and at least one secret key from the plurality of keys is not delivered to the communications unit; and a signature generator for generating an authorization signal from both the at least one secret key and from an authorization message, wherein the authorization message is generated by the communications unit using the at least one communication key.

In another aspect, a method for providing authentication of a subscriber using a subscriber identification device is presented, comprising: generating a plurality of keys; transmitting at least one key from the plurality of keys to a communications device communicatively coupled to the subscriber identification device and holding private at least one key from the plurality of keys; generating a signature at the communications device using both the at least one key transmitted to the communications device and a transmission message; transmitting the signature to the subscriber identification device; receiving the signature at the subscriber identification device; generating a primary signature from the received signature; and conveying the primary signature to a communications system.

In another aspect, a method for providing authentication of a subscriber using a subscriber identification device, comprising: generating a plurality of keys; transmitting at least one key from the plurality of keys to a communications device communicatively coupled to the subscriber identification device and holding private at least one key from the plurality of keys; assigning a weight to the transmission message at the communications device in accordance with a relative importance of the transmission message; generating a signature at the communications device using both the at least one key transmitted to the communications device and the transmission message; transmitting the signature to a communications system if the assigned weight to the transmission message indicates that the transmission message is unimportant; and

transmitting the signature to the subscriber identification device if the assigned weight to the transmission message indicates that the transmission message is important, whereupon the subscriber identification device generates a primary signature from the received signature signal, and then conveys the primary signature to a communications system.

Detailed Description of the Drawings

FIG. 1 is a diagram of an exemplary data communication system.

FIG. 2 is a diagram of a communication exchange between components in a wireless communication system.

FIG. 3 is a diagram of an embodiment wherein a subscriber identification token provides encryption support to a mobile unit.

Detailed Description of the Embodiments

As illustrated in FIG. 1, a wireless communication network 10 generally includes a plurality of mobile stations (also called subscriber units or user equipment) 12a-12d, a plurality of base stations (also called base station transceivers (BTSs) or Node B) 14a-14c, a base station controller (BSC) (also called radio network controller or packet control function 16), a mobile switching center (MSC) or switch 24, a packet data serving node (PDSN) or internetworking function (IWF) 20, a public switched telephone network (PSTN) 22 (typically a telephone company), and an Internet Protocol (IP) network 18 (typically the Internet). For purposes of simplicity, four mobile stations 12a-12d, three base stations 14a-14c, one BSC 16, one MSC 18, and one PDSN 20 are shown. It would be understood by those skilled in the art that there could be any number of mobile stations 12, base stations 14, BSCs 16, MSCs 18, and PDSNs 20.

In one embodiment the wireless communication network 10 is a packet data services network. The mobile stations 12a-12d may be any of a number of different types of wireless communication device such as a portable phone, a cellular telephone that is connected to a laptop computer running IP-based, Web-

browser applications, a cellular telephone with associated hands-free car kits, a personal data assistant (PDA) running IP-based, Web-browser applications, a wireless communication module incorporated into a portable computer, or a fixed location communication module such as might be found in a wireless local loop or meter reading system. In the most general embodiment, mobile stations may be any type of communication unit.

The mobile stations 12a-12d may be configured to perform one or more wireless packet data protocols such as described in, for example, the EIA/TIA/IS-707 standard. In a particular embodiment, the mobile stations 12a-12d generate IP packets destined for the IP network 24 and encapsulate the IP packets into frames using a point-to-point protocol (PPP).

In one embodiment the IP network 24 is coupled to the PDSN 20, the PDSN 20 is coupled to the MSC 18, the MSC 18 is coupled to the BSC 16 and the PSTN 22, and the BSC 16 is coupled to the base stations 14a-14c via wirelines configured for transmission of voice and/or data packets in accordance with any of several known protocols including, e.g., E1, T1, Asynchronous Transfer Mode (ATM), IP, Frame Relay, HDSL, ADSL, or xDSL. In an alternate embodiment, the BSC 16 is coupled directly to the PDSN 20, and the MSC 18 is not coupled to the PDSN 20. In another embodiment of the invention, the mobile stations 12a-12d communicate with the base stations 14a-14c over an RF interface defined in the 3rd Generation Partnership Project 2 "3GPP2", "Physical Layer Standard for cdma2000 Spread Spectrum Systems," 3GPP2 Document No. C.P0002-A, TIA PN-4694, to be published as TIA/EIA/IS-2000-2-A, (Draft, edit version 30) (Nov. 19, 1999), which is fully incorporated herein by reference.

During typical operation of the wireless communication network 10, the base stations 14a-14c receive and demodulate sets of reverse-link signals from various mobile stations 12a-12d engaged in telephone calls, Web browsing, or other data communications. Each reverse-link signal received by a given base station 14a-14c is processed within that base station 14a-14c. Each base station 14a-14c may communicate with a plurality of mobile stations 12a-12d by modulating and transmitting sets of forward-link signals to the mobile stations

12a-12d. For example, as shown in FIG. 1, the base station 14a communicates with first and second mobile stations 12a, 12b simultaneously, and the base station 14c communicates with third and fourth mobile stations 12c, 12d simultaneously. The resulting packets are forwarded to the BSC 16, which provides call resource allocation and mobility management functionality including the orchestration of soft handoffs of a call for a particular mobile station 12a-12d from one base station 14a-14c to another base station 14a-14c. For example, a mobile station 12c is communicating with two base stations 14b, 14c simultaneously. Eventually, when the mobile station 12c moves far enough away from one of the base stations 14c, the call will be handed off to the other base station 14b.

If the transmission is a conventional telephone call, the BSC 16 will route the received data to the MSC 18, which provides additional routing services for interface with the PSTN 22. If the transmission is a packet-based transmission such as a data call destined for the IP network 24, the MSC 18 will route the data packets to the PDSN 20, which will send the packets to the IP network 24. Alternatively, the BSC 16 will route the packets directly to the PDSN 20, which sends the packets to the IP network 24.

FIG. 2 illustrates a method for authenticating a subscriber using a mobile phone in a wireless communication system. A subscriber travelling outside of the range of his or her Home System (HS) 200 uses a mobile unit 220 in a Visited System (VS) 210. The subscriber uses the mobile unit 220 by inserting a subscriber identification token. Such a subscriber identification token is configured to generate cryptographic and authentication information that allows a subscriber to access account services without the need for establishing a new account with the visited system. A request is sent from the mobile unit 220 to the VS 210 for service (not shown in figure). VS 210 contacts HS 200 to determine service to the subscriber (not shown in figure).

HS 200 generates a random number 240 and an expected response (XRES) 270 based on knowledge of the private information held on the subscriber identification token. The random number 240 is to be used as a

challenge, wherein the targeted recipient uses the random number 240 and private knowledge to generate a confirmation response that matches the expected response 270. The random number 240 and the XRES 270 are transmitted from the HS 200 to the VS 210. Other information is also transmitted, but is not relevant herein (not shown in figure). Communication between the HS 200 and the VS 210 is facilitated in the manner described in Fig. 1. The VS 210 transmits the random number 240 to the mobile unit 220 and awaits the transmission of a confirmation message 260 from the mobile unit 220. The confirmation message 260 and the XRES 270 are compared at a compare element 280 at the VS 210. If the confirmation message 260 and XRES 270 match, the VS 210 proceeds to provide service to the mobile unit 220.

Mobile unit 220 sends the random number 240 to the subscriber identification token 230 that has been inserted inside the mobile unit 220 by the subscriber. A Secure Key 300 is stored on the subscriber identification token 230. Both the Secure Key 300 and the random number 240 are used by a key generator 250 to generate the confirmation message 260, a cryptographic Cipher Key (CK) 290, and an Integrity Key (IK) 310. The CK 290 and IK 310 are conveyed to the mobile unit 220.

At the mobile unit 220, the CK 290 can be used to encrypt communications between the mobile unit 220 and the VS 210, so that communications can be decrypted only by the intended recipient of the message. Techniques for using a cryptographic key to encrypt communications are described in co-pending U.S. Patent Application 09/143,441, filed on August 28, 1998, entitled, "Method and Apparatus for Generating Encryption Stream Ciphers," assigned to the assignee of the present invention, and incorporated by reference herein. It should be noted that other encryption techniques can be used without affecting the scope of the embodiments described herein.

The IK 310 can be used to generate a message authentication code (MAC), wherein the MAC is appended to a transmission message frame in order to verify that the transmission message frame originated from a particular party and to verify that the message was not altered during transmission. Techniques

for generating MACs are described in co-pending U.S. Patent Application No. 09/371,147, filed on August 9, 1999, entitled, "Method and Apparatus for Generating a Message Authentication Code," assigned to the assignee of the present invention and incorporated by reference herein. It should be noted that other techniques for generating authentication codes may be used without affecting the scope of the embodiments described herein.

Alternatively, the IK 310 can be used to generate an authentication signature 340 based on particular information that is transmitted separately or with the transmission message. Techniques for generating an authentication signature are described in U.S. Patent 5,943,615, entitled, "Method and Apparatus for Providing Authentication Security in a Wireless Communication System," assigned to the assignee of the present invention and incorporated by reference herein. The authentication signature 340 is the output of a hashing element 330 that combines the IK 310 with a message 350 from the mobile unit 220. The authentication signature 340 and the message 350 are transmitted over the air to the VS 210.

As seen in FIG. 2, the cryptographic key 290 and the integrity key 310are transmitted from the subscriber identification token 230 to the mobile unit 220, which proceeds to generate data frames for public dissemination over the air. While this technique may prevent an eavesdropper from determining the values of such keys over the air, this technique does not provide protection from attack by a rogue shell. A rogue shell can be programmed to accept the CK 290 and the IK 310, and to then store the keys rather than purging the presence of such keys from local memory. Another method to steal keys is to program the mobile unit 220 to transmit received keys to another location. The CK 290 and the IK 310 can then be used to fraudulently bill unauthorized communications to the subscriber. This rogue shell attack is particularly effective in systems wherein the random number generated at the Home System 200 is used in a manner that is insecure, such as the case when the same generated keys are used for an extended period of time.

An embodiment that protects against a rogue shell attack uses the processors and memory in the subscriber identification token to generate an electronic signature that cannot be reproduced by a mobile unit without the insertion of the subscriber identification token.

FIG. 3 illustrates an embodiment for performing local authentication of a subscriber in a wireless communication system. In this embodiment, the subscriber identification token 230 is programmed to generate an authentication response based on a key that is not passed to the mobile unit 220. Hence, if the mobile unit used by a subscriber is a rogue shell, the rogue shell cannot recreate the appropriate authentication responses.

Similar to the method described in FIG. 2, the mobile unit 220 generates a signature signal based upon an IK 310 that is received from the subscriber identification token 230 and a message that is to be sent to the VS 210. However, in the exemplary embodiment, the signature signal is not passed to the VS. The signature signal is passed to the subscriber identification token 230, and is used along with an additional key to generate a primary signature signal. The primary signature signal is sent out to the mobile unit 220, which in turn transmits the primary signature signal to the VS 210 for authentication purposes.

HS 200 generates a random number 240 and an expected response (XRES) 270 based on knowledge of the private information held on the subscriber identification token 230. The random number 240 and the XRES 270 are transmitted to the VS 210. Communication between the HS 200 and the VS 210 is facilitated in the manner described in Fig. 1. The VS 210 transmits the random number 240 to the mobile unit 220 and awaits the transmission of a confirmation message 260 from the mobile unit 220. The confirmation message 260 and the XRES 270 are compared at a compare element 280 at the VS 210. If the confirmation message 260 and the XRES 270 match, the VS 210 proceeds to provide service to the mobile unit 220.

Mobile unit 220 conveys the random number 240 to the subscriber identification token 230 that has been electronically coupled with the mobile unit 220 by the subscriber. A Secure Key 300 is stored on the subscriber

identification token 230. Both the Secure Key 300 and the random number 240 are used by a key generator 250 to generate the confirmation message 260, a Cryptographic Key (CK) 290, an Integrity Key (IK) 310, and a UIM Authentication Key (UAK) 320. The CK 290 and IK 310 are conveyed to the mobile unit 220.

At the mobile unit 220, the CK 290 is used for encrypting transmission data frames (not shown in FIG. 3). The IK 310 is used to generate a signature signal 340. The signature signal 340 is the output of a signature generator 330 that uses an encryption operation or a one-way operation, such as a hashing function, upon the IK 310 and a message 350 from the mobile unit 220. The signature signal 340 is transmitted to the subscriber identification token 230. At the subscriber identification token 230, the signature signal 340 and the UAK 320 are manipulated by a signature generator 360 to generate a primary signature signal 370. The primary signature signal 370 is transmitted to the mobile unit 220 and to the VS 210, where a verification element 380 authenticates the identity of the subscriber. The verification element 380 can accomplish the verification by regenerating the signature signal 340 and the primary signature signal 370. Alternatively, the verification element 380 can receive the signature signal 340 from the mobile unit 220 and only regenerate the primary signature signal 370.

The regeneration of the signature signal 340 and the primary signature signal 370 at the VS 210 can be accomplished by a variety of techniques. In one embodiment, the verification element 380 can receive a UAK 390 and an integrity key from the Home System 200. When the verification element 380 also receives the message 350 from the mobile unit 220, the signature signal can be generated and then used to generate the primary signature element.

The signature generator 360 within the subscriber identification token 230 can comprise a memory and a processor, wherein the processor can be configured to manipulate inputs using a variety of techniques. These techniques can take the form of encryption techniques, hashing functions, or any nonreversible operation. As an example, one technique that can be implemented by the subscriber identification token is the Secure Hash Algorithm (SHA),

promulgated in Federal Information Processing Standard (FIPS) PUB 186, "Digital Signature Standard," May 1994. Another technique that can be performed by the subscriber identification token is the Data Encryption Standard (DES), promulgated in FIPS PUB 46, January 1977. The use of the term "encryption" as used herein does not necessarily imply that operations must be reversible. The operations may be non-reversible in the embodiments described herein.

The key generator 250 can also comprise a memory and a processor. Indeed, in one embodiment, a single processor can be configured to accomplish the functions of the signature generator 360 and the key generator 250. Verification can be performed by calculating the same result from the same inputs at the verification element 380, and comparing the calculated and transmitted values.

A subscriber identification token used in a CDMA system or a GSM system, also known as an R-UIM or a USIM, respectively, can be configured to generate the primary signature signal 370 in the manner described above, i.e., all messages generated by the mobile unit are encrypted and authenticated. However, since the central processing unit in such tokens can be limited, it may be desirable to implement an alternative embodiment, wherein a weight of importance is assigned to a message frame so that only important messages are securely encrypted and authenticated. For example, a message frame containing billing information has more need for increased security than a message frame containing a voice payload. Hence, the mobile unit can assign a greater weight of importance to the billing information message frame and a lesser weight of importance to the voice message frame. When the subscriber identification token receives the signature signals generated from these weighted messages, the CPU can assess the different weights of importance attached to each signature signal and determine a primary signature signal for only the heavily weighted signature signals. Alternatively, the mobile unit can be programmed to convey only the "important" signature signals to the subscriber identification token. This method of selective primary signature signal generation

increases the efficiency of the subscriber identification token by lightening the processing load of the subscriber identification token.

The embodiments described above prevent unauthorized use of a subscriber's account by requiring a more secure transaction between the subscriber identification token and the mobile unit. Since the mobile unit cannot generate a primary signature signal without knowledge of the secret UAK, the mobile unit that is programmed to act as a rogue shell cannot misappropriate subscriber information for wrongful purposes.

The embodiments described above also maximize the processing capability of the subscriber identification token by operating on a signature signal, rather than a message. Typically, a signature signal will have a shorter bit length than a message. Hence, less time is required for the signature generator in the subscriber identification to operate on a signature signal rather than a transmission message frame. As mentioned above, the processing capability of the subscriber identification token is usually much less than the processing capability of the mobile unit. Hence the implementation of this embodiment would provide secure authentication of messages without sacrificing speed.

However, it should be noted that improvements in processor architectures occur at an almost exponential pace. Such improvements consist of faster processing times and smaller processor sizes. Hence, another embodiment for providing local authentication can be implemented wherein the primary signature signal can be generated directly from a message, rather than indirectly through a short signature signal. A mobile unit can be configured to pass a message directly to the subscriber identification token, one with the capability to generate a primary signature signal quickly, rather than passing the message to a signature generating element within the mobile unit. In another embodiment, only a limited number of messages need be passed directly to the subscriber identification token, in accordance with the degree of security needed for said messages.

It should be noted that while the various embodiments have been described in the context of a wireless communication system, the various

embodiments can be further used to provide secure local authentication of any party using an unfamiliar terminal connected in a communications network.

Thus, novel and improved methods and apparatus for performing local authentication of a subscriber in a communication system have been described. Those of skill in the art would understand that the various illustrative logical blocks, modules, circuits, and algorithm steps described in connection with the embodiments disclosed herein may be implemented as electronic hardware, software, firmware, or combinations thereof. The various illustrative components, blocks, modules, circuits, and steps have been described generally in terms of their functionality. Whether the functionality is implemented as hardware, software, or firmware depends upon the particular application and design constraints imposed on the overall system. Skilled artisans recognize the interchangeability of hardware, software, and firmware under these circumstances, and how best to implement the described functionality for each particular application.

Implementation of various illustrative logical blocks, modules, circuits, and algorithm steps described in connection with the embodiments disclosed herein may be implemented or performed with a digital signal processor (DSP), an application specific integrated circuit (ASIC), a field programmable gate array (FPGA) or other programmable logic device, discrete gate or transistor logic, discrete hardware components. A processor executing a set of firmware instructions, any conventional programmable software module and a processor, or any combination thereof can be designed to perform the functions described herein. The processor may advantageously be a microprocessor, but in the alternative, the processor may be any conventional processor, controller, microcontroller, or state machine. The software module could reside in RAM memory, flash memory, ROM memory, EPROM memory, EEPROM memory, registers, hard disk, a removable disk, a CD-ROM, or any other form of storage medium known in the art. An exemplary processor is coupled to the storage medium so as to read information from, and write information to, the storage medium. In the alternative, the storage medium may reside in an ASIC. The

ASIC may reside in a telephone or other user terminal. In the alternative, the processor and the storage medium may reside in a telephone or other user terminal. The processor may be implemented as a combination of a DSP and a microprocessor, or as two microprocessors in conjunction with a DSP core, etc. Those of skill would further appreciate that the data, instructions, commands, information, signals, bits, symbols, and chips that may be referenced throughout the above description are represented by voltages, currents, electromagnetic waves, magnetic fields or particles, optical fields or particles, or any combination thereof.

Various embodiments of the present invention have thus been shown and described. It would be apparent to one of ordinary skill in the art, however, that numerous alterations may be made to the embodiments herein disclosed without departing from the spirit or scope of the invention.

WE CLAIM: